针对无约束最优化问题，通常做法就是对$f(X)$求导，并令$\frac{\partial}{\partial X}f(X)=0$，求解可以得到最优值。如果$f(X)$为凸函数，则可以保证结果为全局最优解。

针对有等式约束的最优化问题，采用**拉格朗日乘数法（Lagrange Multiplier）**[2]进行求解：通过拉格朗日系数$A=[a_1,a_2\dots a_n]^T\in\mathbb{R}^n$把等式约束和目标函数组合成为一个式子，对该式子进行最优化求解：

$$X=\underset{X}{argmin}[f(X)+A^TH(X)]$$

其中，$H(X)=[\,h_1(X),\,h_2(X)\dots\,h_n(X)]^T\in\mathbb{R}^n$，相当于将有等式约束的最优化问题转换成了无约束最优化求解问题，解决方法依旧是对$f(X)+A^TH(X)$的各个参数（X,A）求偏导，并令其等于 0，联立等式求解。

针对有不等式约束的最优化问题，常用的方法是 **KKT 条件（Karush-Kuhn-Tucker Conditions）**[3]：同样地，把所有的不等式约束、等式约束和目标函数全部写为一个式子：

$$L(X,A,B)=f(X)+A^TH(X)+B^TG(X)$$

KKT 条件是说最优值必须满足以下条件：

$$\frac{\partial}{\partial X}L(X,A,B)=0$$

$$H(X)=0$$

$$B^TG(X)=0$$

其中，$B=[b_1,b_2\dots b_m]^T\in\mathbb{R}^m$，$G(X)=[\,g_1(X),\,g_2(X)\dots\,g_m(X)]^T\in\mathbb{R}^m$。KKT 条件是求解最优值$X^*$的必要条件，要使其成为充分必要条件，还需要$f(X)$为凸函数才行。

在 KKT 条件中，$B^TG(X)=0$这个条件最有趣，因为$g_l(X)\le 0$，如果要满足这个等式，需要$b_l=0$或者$g_l(X)=0$。在我们后面的推导中会用到这个性质。

注：

- 无约束优化：直接对$f(x)$求导；
- 有等式约束的最优化问题：采用**拉格朗日乘数法(Lagrange Multiplier)**进行求解；
- 针对有不等式约束约束的最优化问题，采用KKT条件(Karush-Kuhn-Trcker Conditions)；